Am. J. Hum. Genet. 77:1044–1060, 2005

Replication of Putative Candidate-Gene Associations with Rheumatoid Arthritis in >4,000 Samples from North America and Sweden: Association of Susceptibility with *PTPN22, CTLA4,* and *PADI4*

Robert M. Plenge,[2,5] Leonid Padyukov,[9] Elaine F. Remmers,[11] Shaun Purcell,[7] Annette T. Lee,[12] Elizabeth W. Karlson,[5] Frederick Wolfe,[13] Daniel L. Kastner,[11] Lars Alfredsson,[10] David Altshuler,[1,3,4,8] Peter K. Gregersen,[12] Lars Klareskog,[9] and John D. Rioux[2,6,14]

[1]Program in Medical and Population Genetics, [2]Broad Institute of MIT and Harvard, Cambridge, MA; Departments of [3]Genetics and [4]Medicine, Harvard Medical School, [5]Division of Rheumatology, Immunology and Allergy and [6]Department of Neurology, Brigham and Women's Hospital, and [7]Psychiatric and Neurodevelopmental Genetics Unit and [8]Center for Human Genetic Research, Department of Molecular Biology and Diabetes Unit, Massachusetts General Hospital, Boston; [9]Rheumatology Unit and [10]Institute of Environmental Medicine, Karolinska Institutet, Stockholm; [11]Genetics and Genomics Branch, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD; [12]Robert S. Boas Center for Genomics and Human Genetics, North Shore LIJ Institute for Medical Research, Manhasset, NY; [13]National Data Bank for Rheumatic Diseases, Wichita, Kansas; and [14]Université de Montréal and the Montreal Heart Institute/Institut de Cardiologie de Montréal, Montreal

Candidate-gene association studies in rheumatoid arthritis (RA) have lead to encouraging yet apparently inconsistent results. One explanation for the inconsistency is insufficient power to detect modest effects in the context of a low prior probability of a true effect. To overcome this limitation, we selected alleles with an increased probability of a disease association, on the basis of a review of the literature on RA and other autoimmune diseases, and tested them for association with RA susceptibility in a sample collection powered to detect modest genetic effects. We tested 17 alleles from 14 genes in 2,370 RA cases and 1,757 controls from the North American Rheumatoid Arthritis Consortium (NARAC) and the Swedish Epidemiological Investigation of Rheumatoid Arthritis (EIRA) collections. We found strong evidence of an association of *PTPN22* with the development of anti-citrulline antibody–positive RA (odds ratio [OR] 1.49; $P = .00002$), using previously untested EIRA samples. We provide support for an association of *CTLA4* (CT60 allele, OR 1.23; $P = .001$) and *PADI4* (PADI4_94, OR 1.24; $P = .001$) with the development of RA, but only in the NARAC cohort. The *CTLA4* association is stronger in patients with RA from both cohorts who are seropositive for anti-citrulline antibodies ($P = .0006$). Exploration of our data set with clinically relevant subsets of RA reveals that *PTPN22* is associated with an earlier age at disease onset ($P = .004$) and that *PTPN22* has a stronger effect in males than in females ($P = .03$). A meta-analysis failed to demonstrate an association of the remaining alleles with RA susceptibility, suggesting that the previously published associations may represent false-positive results. Given the strong statistical power to replicate a true-positive association in this study, our results provide support for *PTPN22, CTLA4,* and *PADI4* as RA susceptibility genes and demonstrate novel associations with clinically relevant subsets of RA.

Introduction

Rheumatoid arthritis (RA [MIM 180300]) is a common autoimmune disease characterized by chronic, destructive, and debilitating arthritis (Gabriel et al. 1999). The etiology of RA is unknown, but it is thought to have both a genetic and an environmental basis (Firestein 2003). Autoantibodies are detected in ~2/3 of patients with RA and predict more-severe disease (van Zeben et al. 1992; van Jaarsveld et al. 1999; Kroot et al. 2000). The two major types of autoantibodies used clinically to create RA subsets are rheumatoid factor (RF), which is an immunoglobulin specific to the Fc region of IgG, and anti–cyclic citrullinated peptide (CCP) antibodies, which are antibodies directed against peptides in which arginine has been posttranslationally modified to become citrulline (Schellekens et al. 2000). These autoantibodies are strongly correlated but may represent distinct clinical subsets of RA. For decades, it has been known that *HLA-DR* alleles within the major histocompatibility complex (MHC) are associated with RA (Stastny and Fink 1977) and are likely to contribute ~1/3 of the genetic risk in RA (Deighton et al. 1989; Rigby et al. 1991). Although the MHC associations with

Received June 6, 2005; accepted for publication October 3, 2005; electronically published November 1, 2005.

Address for correspondence and reprints: Dr. John D. Rioux, The Broad Institute of MIT and Harvard, Université de Montréal and the Montreal Heart Institute/Institut de Cardiologie de Montréal, 5000 Belanger, Montreal, Canada, H1T 1C8. E-mail: rioux@broad.mit.edu

0002-9297/2005/7706-0015$15.00

RA are complex (Jawaheer et al. 2002; Newton et al. 2004), the majority of the genetic signal from the MHC is explained by multiple alleles at the *HLA-DRB1* locus (MIM 142857) (Hall et al. 1996; Jawaheer et al. 2002). These alleles are known collectively as "shared epitope" (SE) alleles because of their sequence similarity within the third hypervariable region (amino acids 70–74: Leu-x-x-Glu-Arg/Lys [Gregersen et al. 1987]). Non-MHC genes are likely to contribute to RA pathogenesis and disease heterogeneity but have been much more difficult to identify.

Genetic association studies of RA applied to non-MHC candidate genes have led to encouraging yet apparently inconsistent results (see Genetic Association Database Web site). One explanation for the inconsistency is insufficient power to detect modest effects in the context of a low prior probability of a true effect and the failure to assign a stringent *P* value to interpret significance (Hirschhorn and Daly 2005; Wang et al. 2005). Thousands of patients are required to detect an allele with an odds ratio (OR) less than that observed for *HLA-DRB1,* and the vast majority of association studies of RA have been performed with <1,000 samples (notable exceptions include studies by Barton et al. [2004*b*], Begovich et al. [2004], Prokunina et al. [2004], Kochi et al. [2005], and Swanberg et al. [2005]). The inconsistent results and lack of replication make interpreting the significance of these association studies a challenge, and it is difficult to determine which results are true positives and which are false positives.

In the current study, we sought to test alleles with a prior putative association with RA or other autoimmune disorders in a collection of 4,127 samples (2,370 RA cases and 1,757 controls) powered to detect modest genetic effects (Hirschhorn et al. 2002; Lohmueller et al. 2003). Our goal was to determine which alleles replicate in our collection from Sweden and North America, thus providing evidence of a true-positive association. We first performed a thorough review of the literature on RA candidate-gene association studies, to identify alleles with suggestive (but not definitive) evidence of association with the development of RA. We also reviewed the literature to identify alleles with convincing evidence of association with autoimmune diseases other than RA, since we hypothesized that common pathogenic mechanisms may be shared among RA and other autoimmune diseases (Yamada and Ymamoto 2005). Here, we describe our results of testing 17 alleles in 14 genes for association with RA susceptibility in two independent clinical collections consisting of Swedish or North American individuals of European descent. Our results validate the hypothesis that the R620W variant of *PTPN22* (MIM 600716) is associated with the development of RF-positive (RF+) and anti-CCP–positive (CCP+) RA and provide support for an association of RA with variants in *PADI4* (MIM 605347) and *CTLA4* (MIM 123890).

Material and Methods

Clinical Samples

Two case-control clinical collections were used to test all 17 alleles for association between genetic variation and the development of RA. Clinical characteristics for both collections are provided in table 1. All individuals are of European descent. In total, 4,127 samples (2,370 RA cases and 1,757 controls) were available for analysis. This study was approved by the MIT Broad Institute Institutional Review Board.

The first cohort, the Epidemiological Investigation of Rheumatoid Arthritis (EIRA), is an inception cohort collected in Sweden between May 1996 and February 2001. This cohort has been described in more detail elsewhere (Stolt et al. 2003; Padyukov et al. 2004). Briefly, a case is defined as an individual who fulfills the American College of Rheumatology (ACR) 1987 criteria for the classification of RA (Arnett et al. 1988) and who has never before been diagnosed with RA. For each potential case, a control subject was randomly selected from the study base, with the subject's age, sex, and residential area taken into consideration. In total, 1,530 confirmed RA cases and 881 controls were available for analysis. Low-resolution HLA typing was performed on ~94% of the samples by use of the sequence-specific primer-PCR method (HLA-DR low-resolution kit) as described elsewhere (Olerup and Zetterquist 1992).

The second collection is the North American Rheumatoid Arthritis Consortium (NARAC). This collection of affected sibling pairs has been described in more detail elsewhere (Jawaheer et al. 2001, 2004; North American Rheumatoid Arthritis Consortium Web site). Briefly, a case is defined as an individual who fulfills the ACR 1987 criteria for RA and who has at least one sibling with RA. At least one sibling must have documented erosions on hand radiographs, and at least one sibling must have disease onset between the ages of 18 and 60 years. In total, 840 affected cases from 463 families were used in our analysis. Unrelated matched controls (n = 867) were selected from 20,000 healthy individuals who were enrolled in the New York Cancer Project (NYCP) (see New York Cancer Project Web site). Two controls were matched to a single randomly chosen affected sibling on the basis of sex, age (birth decade), and ethnicity (grandparental country/region of origin). Complete data for *PTPN22* (Begovich et al. 2004; Lee et al. 2005) and partial data for *CARD15* (Addo et al. 2005) have been published elsewhere for this cohort. *PADI4* genotype data were generated at the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) by use of an overlapping set of NARAC and NYCP samples:

Table 1

Clinical Characteristics of EIRA and NARAC Cohorts

Characteristic	EIRA	NARAC
No. of cases	1,530	840
No. of controls	881	876
Study design	Unrelated cases with matched unrelated controls	Affected sib cases with matched unrelated controls
Ascertainment	Incident cases (May 1996–February 2001)	Radiographic erosions and 1 sibling aged 18–60 years
Geographic ancestry	European	European
Country of origin	Sweden	United States
Female sex (%)	70.6	77.4
Age at onset (mean ± SD years)	51.7 ± 12.8	39.4 ± 13.1
Disease duration (mean ± SD years)	<3	16.2 ± 11.8
Autoantibody positivity (%):		
RF+	66.2	78.2
CCP+	60.9	77.8
RF+ or CCP+	72.5	84.7
HLA SE status (allele frequency):		
Any	.49	.62
SE subtype:		
*HLA-DRB1*04*	.34	.48
*HLA-DRB1*01*	.13	.14
*HLA-DRB1*10*	.01	.03
Clinical features (%):		
Radiographic erosions	Unknown	95.3
Thyroiditis (self-reported)	Unknown	19

NOTE.—Details on the EIRA cohort (Stolt et al. 2003; Padyukov et al. 2004) and NARAC cohort (Jawaheer et al. 2001, 2004) have been published elsewhere.

601 NARAC cases and NYCP controls were common to the current data set, and 1,115 individuals were unique.

In addition to these two cohorts, an additional collection ("arthritis cohort 3") of 1,417 patients with RA (with no matched controls) was used to investigate the influence of *PTPN22* alleles on RA phenotype in males and females and to study age-at-onset effects. These samples were used in a case-only analysis comparing allele frequencies in autoantibody-positive (RF+ and/or CCP+) disease with those in autoantibody-negative (RF− and/or CCP−) disease. Consequently, no matched controls were necessary. Patient samples were obtained from the Wichita Rheumatic Disease Data Bank (Wolfe et al. 2003); The Arthritis, Rheumatism and Aging Medical Information System; the National Inception Cohort of Rheumatoid Arthritis Patients (Fries et al. 2002) (130 RF+ male cases, 401 RF+ female cases, 42 RF− male cases, and 139 RF− female cases); and the Study Of New Onset Rheumatoid Arthritis (SONORA) (Bombardier et al. 2002; S.O.N.O.R.A. Web site) (128 RF+ male cases, 313 RF+ female cases, 70 RF− male cases, and 194 RF− female cases). These RA sample sets were combined, and, for comparison with the EIRA and NARAC collections, only patients of European descent were analyzed (n = 1,417 total RA cases).

Literature Review

We reviewed the literature published in English before December 31, 2004, to identify non-MHC alleles that had results suggesting a putative association with RA. Inclusion criteria for a putative susceptibility allele were (*a*) location within a non-MHC gene; (*b*) association with RA susceptibility—that is, the presence or absence of RA (we did not include an allele when the reported association was limited to a secondary phenotype, such as disease severity, because our primary purpose was to replicate previous associations with RA susceptibility); (*c*) a SNP or insertion-deletion (indel) polymorphism (i.e., no microsatellite markers); (*d*) genotyping in at least 200 patients with RA, with samples other than EIRA or NARAC; and (*e*) significance at a level <.001 in one study or <.05 in two or more studies (nominal P values based on allele frequencies in cases vs. controls). We identified >200 published association studies of RA. From these studies, we identified 10 alleles within nine genes that met our inclusion criteria for testing in our combined RA sample collection (table 2). Specific alleles at *PADI4* (Suzuki et al. 2003), *CTLA4* (Rodriguez et al. 2002; Vaidya et al. 2002), *IL4* (Cantagrel et al. 1999; Maksymowych et al. 2002), *SLC22A4* (Tokuhiro et al. 2003), *MIF* (Barton et al. 2003), and *TNFRSF1B* (Barton et al. 2001; Dieude et al. 2002; Kyogoku et al. 2003)

have been associated with the development of RA in some studies but not in others (Dahlqvist et al. 2002; Fabris et al. 2002; Huang et al. 2002; Yen et al. 2003; Barton et al. 2004*b*; Miterski et al. 2004; Orozco et al. 2004; van der Helm-van Mil et al. 2004; Newman et al. 2005). Other genes (e.g., *IL3* [Yamada et al. 2001], *RUNX1* [Tokuhiro et al. 2003], and *HAVCR1* [Chae et al. 2004]) have demonstrated a positive association in a single study but have not yet been tested in an independent collection for replication. Only *PTPN22* has been convincingly replicated in more than one independent study (Begovich et al. 2004; Hinks et al. 2005; Lee et al. 2005; Orozco et al. 2005; Steer et al. 2005; Viken et al. 2005; Zhernakova et al. 2005), including the NARAC samples.

Our replication study was limited to those alleles with the strongest evidence of a prior association, for efficiency and to minimize multiple-hypothesis testing. We chose the CT60 variant (*rs3087243*) of *CTLA4,* rather than the serine→threonine missense SNP at amino acid 17 (*rs231775*), because the CT60 variant clearly influences the risk of Graves disease and type 1 diabetes (Ueda et al. 2003). We chose the slc2F1 (*rs2073838*) allele of *SLC22A4* because this allele had the strongest association with RA in a Japanese population (Tokuhiro et al. 2003). We chose the PADI4_94 allele (*rs2240340*) of *PADI4* because it had the strongest association in a Japanese population (Suzuki et al. 2003). *IL4* did not strictly fit the criteria, but it was included in our study because it is located within the 5q31 cytokine cluster associated with inflammatory bowel disease (Rioux et al. 2001) and is in very close proximity to *SLC22A4* and *IL3,* both of which have reported evidence of association with RA. Although alleles within *FCGR3A* satisfied the inclusion criteria, they were not studied further, because *FCGR3A* was not amenable to genotyping by use of our platform. *PTPN22* has been tested in NARAC, but we included it in our analysis to replicate the result in EIRA and to explore clinically relevant RA subsets and epistatic interactions.

We also hypothesized that an allele associated with another autoimmune disease has an increased prior probability of being associated with RA—compared with a random allele chosen from the genome—on the basis of shared pathogenic mechanisms among autoimmune diseases. We selected seven alleles from five additional genes with convincing evidence of causality in either inflammatory bowel disease (*CARD15* [Hugot et al. 2001; Ogura et al. 2001], *DLG5* [Stoll et al. 2004], and *IBD5* [Rioux et al. 2001; Daly and Rioux 2004]) or type 1 diabetes (*SUMO4* [Guo et al. 2004] and *CTLA4* [Ueda et al. 2003]), although previous studies have not shown an association between *CARD15* and RA (Ferreiros-Vidal et al. 2003; Steer et al. 2003; Addo et al. 2005).

Genotyping

Genotyping was performed, as described elsewhere (Gabriel et al. 2002), by primer extension of multiplex products with detection by matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy (Tang et al. 1999) with the use of a Sequenom platform. All genotyping was performed at the Broad Institute, except for the *PADI4* variant in NARAC, which was genotyped at the NIAMS. The average completeness of genotypes was 97.8% at the Broad Institute and 98.7% at the NIAMS. Using genotypes performed in replicate (n = 3,814), we documented a consensus error rate of 0.5% in the analyzed data. All SNPs were in Hardy-Weinberg equilibrium within individual sample collections ($P > .01$).

Statistical Analysis

To test our primary hypothesis (that an allele is associated with the development of RA) and secondary hypotheses (that particular alleles are associated with a related RA phenotype reported previously in the literature) in both EIRA and NARAC samples, one-tailed P values were calculated using 2-by-2 contingency tables of allele frequencies and a Fisher's exact test, and ORs were reported relative to the published "susceptibility" allele. Results from the EIRA and NARAC cohorts were combined by Mantel-Haenszel meta-analysis of the ORs (Lohmueller et al. 2003), because we hypothesized that a true susceptibility allele would have an effect in both populations, despite possible genetic, ethnic, or clinical differences (Ioannidis et al. 2004). A P value of .05 was considered significant for replication of an allele with previous evidence of association (Wacholder et al. 2004). Prior to combining the EIRA and NARAC results, we tested for homogeneity in the ORs, using a Pearson χ^2 goodness-of-fit test as described elsewhere (Lohmueller et al. 2003). The EIRA and NARAC cohorts demonstrated P values for homogeneity >.01 for all loci except *PTPN22* ($P = .0002$). *PADI4* ($P = .03$) and *CTLA4* ($P = .08$) demonstrated a trend toward significant difference. We also tested our primary hypothesis by using logistic regression, controlling for familial relatedness in NARAC, age, *PTPN22* and SE allele status, and autoantibody status. Relatedness among siblings was handled by use of robust (Huber/White/sandwich) variance estimates (Williams 2000). Genotype data from the NARAC cohort for two genes (*PTPN22* [Begovich et al. 2004; Lee et al. 2005] and *CARD15* [Addo et al. 2005]) have been reported elsewhere. The NARAC genotypes are included here for completeness, although tests of significance were based on EIRA genotype data. Genetic models were tested using a log-linear model of sample allele frequencies conditional on disease status (see Genetic Association Models Web site). Power cal-

Table 2

Results Testing Our Primary Hypothesis That an Allele with Prior Evidence of an Association with RA Susceptibility or Another Autoimmune Disease Replicates in Our Collection of 4,127 Clinical Samples from the EIRA and NARAC Clinical Collections

Gene (Variant), dbSNP, and Collection	Allele	Cases				Controls				OR (95% CI)	*P*
		1/1	1/2	2/2	Frequency	1/1	1/2	2/2	Frequency		
PTPN22 (R620W): *rs2476601*:	T										
EIRA		43	389	1,081	.16	16	187	671	.13	1.30 (1.09–1.54)	.0015
NARAC[a]		20	236	568	.17	9	125	713	.08	2.18 (1.76–2.70)	2.15×10^{-13}
Pooled										1.59 (1.39–1.82)	4.8×10^{-12}
PADI4 (PADI4_94): *rs2240340*:	A										
EIRA		244	737	517	.41	153	393	312	.41	1.01 (.89–1.14)	.47
NARAC		184	447	264	.46	134	335	279	.40	1.24 (1.08–1.42)	.001
Pooled										1.10 (1.00–1.21)	.02
TNFRSF1B (R196M): *rs1061622*:	G										
EIRA		84	540	834	.24	61	320	472	.26	.92 (.80–1.05)	.90
NARAC		36	272	455	.23	39	243	397	.24	.94 (.79–1.12)	.77
Pooled										.93 (.83–1.03)	.92
RUNX1 (runx1): *rs2268277*:	C										
EIRA		207	705	599	.37	132	391	347	.38	.97 (.86–1.10)	.68
NARAC		123	382	324	.38	108	393	342	.36	1.08 (.94–1.24)	.16
Pooled										1.02 (.93–1.12)	.36
MIF (−173 G/C): *rs755622*:	C										
EIRA		51	471	992	.19	35	254	585	.19	1.03 (.88–1.19)	.39
NARAC		26	227	579	.17	31	226	590	.17	.98 (.82–1.18)	.59
Pooled										1.01 (.90–1.13)	.46
HAVCR1 (5383_5397del): *rs6149307*:	Non del										
EIRA		33	389	1,081	.15	19	244	611	.16	.93 (.79–1.09)	.83
NARAC		22	232	578	.17	19	202	627	.14	1.21 (1.00–1.46)	.03
Pooled										1.04 (.92–1.17)	.28
HAVCR1 (5509_5511delCAA):	del										
EIRA		48	410	1,026	.17	23	269	578	.18	.93 (.80–1.09)	.83
NARAC		27	246	491	.20	36	247	396	.24	.80 (.67–.95)	.99
Pooled										.87 (.77–.98)	.99
SLC22A4 (slc2F1): *rs2073838*:	A										
EIRA		13	210	1,293	.08	7	107	760	.07	1.13 (.90–1.42)	.15
NARAC		5	131	693	.09	6	128	713	.08	1.03 (.81–1.32)	.91
Pooled										1.09 (.92–1.28)	.16
IL3 (rIL3-16): *rs31480*:	C										
EIRA		837	567	109	.74	505	298	71	.75	.96 (.84–1.10)	.73
NARAC		501	285	40	.78	494	305	48	.76	1.09 (.93–1.29)	.15
Pooled										1.01 (.91–1.12)	.40
IL4 (−590T): *rs2243250*:	T										
EIRA		56	507	940	.21	38	253	584	.19	1.12 (.97–1.30)	.07
NARAC		22	219	594	.16	22	254	571	.18	.88 (.73–1.05)	.93
Pooled										1.01 (.90–1.14)	.40
CTLA4 (CT60): *rs3087243*:	G										
EIRA		595	680	230	.62	337	396	145	.61	1.05 (.93–1.19)	.21
NARAC		308	387	133	.61	254	426	165	.55	1.23 (1.08–1.42)	.001
Pooled										1.13 (1.03–1.24)	.004

(*continued*)

Table 2 (continued)

Gene (Variant), dbSNP, and Collection	Allele	Cases 1/1	Cases 1/2	Cases 2/2	Cases Frequency	Controls 1/1	Controls 1/2	Controls 2/2	Controls Frequency	OR (95% CI)	*P*
CARD15 (R702W): *rs2066844*:	T										
EIRA		1	56	1,452	.02	0	36	833	.02	.93 (.61–1.41)	.68
NARAC[a]		0	71	752	.05	0	66	773	.05	1.10 (.78–1.55)	.32
Pooled										1.03 (.79–1.34)	.42
CARD15 (G908R): *rs2066845*:	C										
EIRA		0	11	1,510	.004	0	9	866	.005	.70 (.29–1.69)	.85
NARAC[a]		0	28	809	.02	0	31	817	.02	.91 (.55–1.53)	.73
Pooled										.86 (.55–1.34)	.75
CARD15 (1007fs or 3020insC): *rs2066847*:	insC										
EIRA		0	31	1,481	.01	0	24	844	.01	.74 (.43–1.26)	.89
NARAC[a]		2	30	805	.02	0	36	812	.02	.96 (.60–1.54)	.85
Pooled										.86 (.60–1.22)	.81
DLG5 (R30Q): *rs1248696*:	A										
EIRA		12	242	1,239	.09	5	148	720	.09	.98 (.80–1.21)	.59
NARAC		7	167	653	.11	8	141	688	.09	1.19 (.95–1.49)	.08
Pooled										1.07 (.92–1.25)	.19
SUMO4 (M55V): *rs237025*:	G										
EIRA		333	729	451	.46	180	458	235	.47	.97 (.86–1.09)	.70
NARAC		178	449	202	.49	263	397	189	.54	.79 (.69–.91)	.99
Pooled										.89 (.81–.97)	.99

Note.—The putative susceptibility allele is listed. Genotype counts in the two cohorts are expressed relative to the putative susceptibility allele, where allele "1" is the putative susceptibility allele, and allele "2" is the nonsusceptible allele. One-tailed *P* values were calculated on the basis of allele frequency, relative to the putative susceptibility allele. *PTPN22* independently replicates in the EIRA cohort ($P = .0015$), and *PADI4* and *CTLA4* replicate in the EIRA and NARAC cohorts combined ($P < .05$). Homogeneity between EIRA and NARAC was calculated using a Pearson χ^2 goodness-of-fit test, and only *PTPN22* demonstrated a *P* value <.01 between the two collections. All alleles demonstrated Hardy-Weinberg *P* values >.01.

[a] NARAC data that have been published previously.

culations were performed under a multiplicative model ($\alpha = 0.01$) by use of the program by Purcell et al. (2003).

Exploratory and stratified analyses for *PTPN22, CTLA4,* and *PADI4* were tested using a Fisher's exact test and logistic regression. Stratified analyses for the association of *PTPN22, CTLA4,* and *PADI4* with RA case/control status were performed for three sets of subgroups: male versus female, RF+ versus RF−, and CCP+ versus CCP−. Subsequently, RA phenotype analyses for the association of *PTPN22, CTLA4,* and *PADI4* within each nonoverlapping subgroup included comparison of male RA cases with female RA cases, comparison of RF+ cases with RF− cases, and comparison of CCP+ cases with CCP− cases. In addition, RF+ and CCP+ cases were compared with all controls. Analysis of sex-specific effects for *PTPN22* was performed by pooling the allele frequencies across all three clinical collections and comparing male RF+ cases with female RF+ cases (by logistic regression). Linear regression was used to test whether a genotype influenced age at disease onset. Epistatic interactions between pairs of alleles were tested among *PTPN22, CTLA4, PADI4,* and *HLA-DRB1* SE alleles, as recently advocated (Marchini et al. 2005), by use of logistic regression to determine whether genotype distributions were statistically different between cases and controls. High-resolution genotyping was available for NARAC, and the SE alleles included *HLA-DRB1*0101, *0401, *0404, *0405, *0408,* and **1001.* In the EIRA cohort, *HLA-DRB1*01, *04,* and **10* were considered SE alleles (Fries et al. 2002), since only low-resolution genotyping was available. On the basis of NARAC data, we estimated that the fraction of *HLA-DRB1*01, *04,* and **10* individuals in the EIRA cohort who carry the SE allele is ~92%, 96%, and 81%, respectively. To estimate the false-positive rate, we implemented a Bonferroni correction of our *PTPN22* results for three hypotheses (age, sex, and epistasis with *HLA*) and of our *CTLA4* and *PADI4* results for five hypotheses (age, sex, autoantibody status, and epistasis with *HLA* or *PTPN22*).

We performed a meta-analysis of all studies with published allele-frequency data for individuals of European

ancestry, together with our genotype data, as described elsewhere (Lohmueller et al. 2003). We first tested for homogeneity in the ORs across studies, using a Pearson χ^2 goodness-of-fit test. When studies across all ethnic populations were analyzed together, we observed heterogeneity for *IL4, TNFRSF1B,* and *MIF.* To minimize heterogeneity, we excluded studies performed in non-European populations. Among studies of individuals of European ancestry, we observed no evidence of significant heterogeneity ($P > .01$). We calculated the pooled OR under a fixed-effects model (Mantel-Haenszel meta-analysis). The original hypothesis-generating publication was excluded from the meta-analysis, when applicable, to avoid bias from the null hypothesis of no effect (Hirschhorn et al. 2002). Genotype data were available for *IL4* (Cantagrel et al. 1999), *MIF* (Miterski et al. 2004), *TNFRSF1B* (Barton et al. 2001; Dahlqvist et al. 2002; Fabris et al. 2002; Kyogoku et al. 2003; Yen et al. 2003; van der Helm-van Mil et al. 2004), *CARD15* (Ferreiros-Vidal et al. 2003; Steer et al. 2003; Addo et al. 2005), and *CTLA4* (Barton et al. 2004*b*; Orozco et al. 2004) but were unavailable from some studies (Dieude et al. 2002; Maksymowych et al. 2002; Newman et al. 2005). We excluded *MIF* data from Barton et al. (2003), since 25% of the cases with inflammatory polyarthritis failed to develop RA. Only a single study (Martinez et al. 2005) has tested the PADI4_94 variant (*rs2240340*)—the variant was not tested by Barton et al. (2004*a*). Across all studies, there were insufficient data on clinically relevant subsets of RA (e.g., subsets based on autoantibody status and sex) available to perform a more refined meta-analysis.

Results

Description of EIRA and NARAC Clinical Cohorts and Estimation of Study Power

To maximize our power to detect a true causal RA allele, we included >4,000 clinical samples from two distinct cohorts: the EIRA (Stolt et al. 2003; Padyukov et al. 2004) and the NARAC (Jawaheer et al. 2001, 2004). The clinical features of the EIRA and NARAC cohorts are shown in table 1. The combined resource of >4,000 total case-control samples provides good power to detect even modest genetic effects; we have >80% power to detect the published OR for 14 of the 17 alleles selected for replication and >50% power to detect an OR $\geqslant$1.25 for all alleles (see the "Material and Methods" section). Thus, failure to replicate in our clinical collection substantially reduces the probability that an allele is a true causal allele.

Selection of Candidate Alleles and Testing of Primary Hypothesis

To maximize further our ability to detect a true positive result, we selected alleles from the published genetic association studies of RA and other autoimmune diseases. Our goal for the literature review was to identify alleles with suggestive evidence of association with RA ($P < .001$ in one study or $P < .05$ in two or more studies) or convincing association with a different autoimmune disease. Our primary hypothesis was that an allele that satisfies these criteria is associated with susceptibility to RA in our collection. Applying our inclusion criteria (see the "Material and Methods" section) to the RA and autoimmune literature, we identified 17 alleles within 14 genes with a putative association with RA or another autoimmune disease (table 2). Ten alleles within nine genes had suggestive statistical evidence in favor of an association with the development of RA. Only the R620W variant of *PTPN22* had been convincingly replicated in more than one study (Hinks et al. 2005; Lee et al. 2005; Orozco et al. 2005; Steer et al. 2005; Viken et al. 2005; Zhernakova et al. 2005). Seven alleles within five genes had convincing evidence in favor of an association with a chronic autoimmune disease (either inflammatory bowel disease or type 1 diabetes), although previous studies have not shown an association between one of those five genes, *CARD15,* and RA (Ferreiros-Vidal et al. 2003; Steer et al. 2003; Addo et al. 2005).

To determine whether any of the 17 alleles identified through our literature review is associated with RA susceptibility (primary hypothesis), we genotyped each allele in our combined clinical collection of >4,000 case-control samples. For each allele, we first analyzed the cohorts individually; second, we performed a formal test for heterogeneity in the OR across the two cohorts; and third, we combined the samples in a meta-analysis to provide the best estimate of the OR (under the assumption of a homogeneous effect across cohorts). Of the 17 alleles tested, 3 replicated association, with a one-tailed P value <.05 (table 2) in the combined samples: *PTPN22*, *PADI4,* and *CTLA4*. *PTPN22* had not been tested previously in the EIRA cohort and thus represents an independent replication. The data for each of the three genes fit best under a multiplicative genetic model (compared with general, dominant, and recessive models). The *CTLA4* result remained significant after we controlled for self-reported thyroiditis (with which *CTLA4* demonstrates a genetic association [Ueda et al. 2003]) in the NARAC cohort (OR 1.19; 95% CI 1.03–1.38; $P = .02$). We did observe a P value <.05 for *SUMO4* and *HAVCR1,* but it was for the opposite allele than the published susceptibility allele. (Because we report our results in table 2 with respect to the published putative susceptibility allele, *SUMO4* and *HAVCR1*

have ORs that are significantly <1.0 but have *P* values close to 1.) None of the other alleles demonstrated a significant association in these two cohorts (table 2), even after we controlled for the effects of *HLA-DRB1* and *PTPN22* (data not shown).

Selection and Testing of Secondary Hypotheses

RA is a heterogeneous disease, and it is possible that an allele is associated with only a clinical subset of RA. Therefore, we used our literature review to define specific secondary hypotheses to test in our RA clinical collection, to ensure that we did not miss a positive association. Each secondary hypothesis was based on data from a previous publication that suggested a specific genetic association. The secondary hypotheses relate either to an RA subphenotype or to a specific genetic model. We defined five secondary hypotheses (with published ORs): (1) *PTPN22* is associated with RF+ but not RF− disease (OR 2.0) (Begovich et al. 2004; Lee et al. 2005); (2) homozygotes for the *PADI4* susceptibility allele are more likely to be CCP+ than are the other genotypes combined (OR 3.25) (Suzuki et al. 2003); (3) individuals homozygous for the putative susceptibility alleles at both *SLC22A4* and *RUNX1* have a marked increase in RA disease risk under a dominant (but not epistatic) model (OR 9.03) (Tokuhiro et al. 2003); (4) *TNFRSF1B* is associated with familial RA but not sporadic RA (OR 1.69) (Barton et al. 2001; Dieude et al. 2002); and (5) *IL3* is associated with RA in females with early-onset disease (OR 1.89) (Yamada et al. 2001). No other secondary hypotheses with published statistical evidence were identified for the remaining 11 alleles. Given the high ORs, we have >80% power to detect each hypothesis, with the exception of the one regarding *SLC22A4* and *RUNX1,* for which there is limited power to detect an association. We modified the hypothesis to ask whether individuals homozygous at *RUNX1* and either heterozygous or homozygous at *SLC22A4* ("susceptible") were at an increased risk of developing RA, compared with all other individuals ("nonsusceptible"). We have >80% power to detect this hypothesis (estimated OR 1.52, on the basis of genotype data from Tokuhiro et al. [2003]).

Each of our five secondary hypotheses was tested in the EIRA and NARAC clinical collections. As shown in table 3, we replicated the finding that *PTPN22* is associated with RF+ disease but not RF− disease, using the previously untested EIRA cohort ($P = .0001$ in EIRA). We extend this observation in autoantibody-positive disease to include an association with CCP+ cases but not CCP− cases (table 3) in the EIRA cohort, which is consistent with previously published clinical observations that these autoantibodies are strongly correlated. Using data combined from EIRA and NARAC, we observe a trend toward an association with *PADI4* and anti-CCP antibody status, as previously suggested, but this was not statistically significant (table 4). We found no significant evidence to support an additive model between *SLC22A4* and *RUNX1* (table 5), although we have limited power to detect the model proposed by Tokuriho et al. (2003), given the allele frequencies observed in our population; we found no significant evidence to support an association between *IL3* and RA in females with early-onset disease (table 6); and we found no significant evidence for an association of *TNFRSF1B* with familial RA, represented by the NARAC affected sibling pairs (table 2).

Exploring Novel Hypotheses for PTPN22

Having established clear evidence for an association of RA susceptibility with *PTPN22* in our samples, we sought to explore whether the R620W variant is associated with a clinically relevant subset of RA. Each exploratory hypothesis was defined on the basis of phenotypes of clinical relevance to RA, as established through epidemiological studies, animal models, or con-

Table 3

Results Testing Our Secondary Hypothesis That *PTPN22* Is Associated with Autoantibody-Positive (AutoAb+) but Not Autoantibody-Negative (AutoAb−) Disease in the EIRA Cohort

EIRA Group	Allele			Cases vs. All Controls		AutoAb+ vs. AutoAb− Cases	
	T	C	MAF	OR (95% CI)	*P*	OR (95% CI)	*P*
RF+ cases	321	1,565	.17	1.43 (1.19–1.72)	.0001	1.37 (1.10–1.72)	.003
RF− cases	126	844	.13	1.04 (.82–1.32)	.73		
CCP+ cases	319	1,499	.18	1.49 (1.23–1.79)	.00002	1.43 (1.16–1.77)	.0004
CCP− cases	151	1,017	.13	1.04 (.83–1.29)	.40		

Note.—The two categories of autoantibodies are RF and CCP. The *PTPN22* allele-frequency distribution was compared among RF+ cases, RF− cases, and all EIRA controls. Only RF+ patients demonstrated a statistically significant difference in allele frequency when compared with all controls ($P = .0001$). The difference between RF+ and RF− cases is statistically significant ($P = .003$). We also stratified EIRA patients on the basis of CCP antibody status. We found that the allele-frequency distribution in CCP+ cases (but not CCP− cases) is statistically different from that in controls ($P = .00002$). The difference between CCP+ and CCP− cases is statistically significant ($P = .0004$).

Table 4

Results Testing Our Secondary Hypothesis That *PADI4* Is Associated with CCP+ but Not CCP− Disease in the Combined EIRA and NARAC Collections

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

vincing genetic studies of the *HLA-DRB1* SE alleles. The hypotheses were considered exploratory because no suggestive genetic association for the individual genes had been reported previously (in contrast to our primary and secondary hypotheses). Our three exploratory hypotheses include: (1) the allele has a greater effect in one sex than in the other, (2) the allele influences age at disease onset, and (3) the allele is epistatic with *HLA-DRB1* SE alleles.

In an analysis of EIRA and NARAC cases, together with a separate case-only collection of 1,417 RA cases ("arthritis cohort 3" [Lee et al. 2005]), we found modest support that the *PTPN22* susceptibility allele has a greater effect in males than in females ($P = .03$) (table 7). This effect was most prominent in the EIRA cohort and the arthritis cohort 3. We also observed an age-at-onset effect for *PTPN22*: carriers of the susceptible T allele develop RA 2 years earlier than do those not carrying the allele (T/T or C/T mean ± SD, 48.2 ± 14.0; C/C, 50.1 ± 14.3; $P = .004$). We found no evidence to support an epistatic interaction between *PTPN22* and SE alleles (data not shown) in the EIRA cohort, which is consistent with previously published reports on NARAC (Begovich et al. 2004). After application of a conservative Bonferroni correction for three hypotheses, the sex effect is no longer significant at a P value <.05, whereas the age-at-onset effect remains significant ($P = .012$).

Stratified Analyses to Explain Heterogeneity for PTPN22, CTLA4, *and* PADI4

We observed a difference in the magnitude of the genetic effect, as measured by the ORs, between the EIRA and NARAC cohorts: *PTPN22* is stronger and more significant in NARAC than in EIRA, and *CTLA4* and *PADI4* are only significant in the NARAC cohort. As part of our initial analysis, we formally tested for heterogeneity between the ORs of the two cohorts, and we observed a clear significant difference for *PTPN22* ($P = .0002$) and a trend toward significance for *PADI4* ($P = .03$) and *CTLA4* ($P = .08$). Because the ascertainment criteria and clinical characteristics between the two cohorts are distinct, we sought to explore whether a clinically relevant RA subset or an epistatic interaction might explain the difference in observed ORs. Therefore, we stratified our analysis by available clinically relevant subsets of RA (sex, autoantibody status, and age at onset) and epistasis with *HLA* and/or *PTPN22*.

Stratified analyses demonstrated that *CTLA4* has a stronger association in patients positive for RF or CCP autoantibodies (table 8). This effect is most apparent in a pooled analysis of EIRA and NARAC, when CCP+ cases are compared with all controls (OR 1.18; 95% CI 1.07–1.31; $P = .0006$). The result remains significant after a Bonferroni correction for five hypotheses ($P = .003$). No other clinically relevant subset demonstrated a significant association, and we found no evidence of epistasis between any two combinations of *CTLA4* with *HLA* or *PTPN22* alleles and *PADI4* with *HLA* or *PTPN22* alleles (data not shown). Overall, however, we were unable to find any variable(s) to explain the OR differences for the three genes between the EIRA and NARAC cohorts.

Meta-Analysis with Published Data

Alleles with modest impact on disease susceptibility may be missed by any single study. We therefore performed a meta-analysis, combining our results with previously published results for individuals of European descent, to determine whether subtle effects across all published studies could be detected, under the hypothesis that a true susceptibility allele has an effect in different populations despite possible ethnic or clinical differences. The results, shown in table 9, fail to demonstrate an association of RA with *CARD15, IL4, MIF,* or *TNFRSF1B*. In contrast, the meta-analysis of previously published data with our data for *CTLA4* (table 10) and *PADI4* (combined OR 1.12; 95% CI 1.03–1.23 [Martinez et al. 2005]) demonstrates continued evidence of association with RA.

Discussion

Genetic association is a powerful statistical methodology to identify variants that influence the risk of developing a common disease such as RA. However, inconsistent results across multiple studies make it difficult to interpret which results represent true-positive association and, therefore, which genes truly influence the risk of RA. One explanation for the inconsistent results is insufficient power to detect modest effects in the context of a low prior probability of a true effect (Hirschhorn and Daly 2005; Wang et al. 2005). To overcome this

Table 5

Results Testing Our Secondary Hypothesis That *SLC22A4* and *RUNX1* Susceptible Genotypes Additively Influence the Risk of RA

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

Table 6

Results Testing Our Secondary Hypothesis That *IL3* Is Associated with Early-Onset RA in Females

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

limitation, we designed a study to test alleles with an increased prior probability of disease association in a sample collection powered to detect modest genetic effects. Our results confirm the association of *PTPN22* with autoantibody-positive RA and provide support for an association with *CTLA4* and *PADI4*. Stratified analysis of our data reveals a novel association with age at onset (*PTPN22*), sex (*PTPN22*), and autoantibody-positive RA (*CTLA4*), results that need to be replicated in an independent data set.

Our results in the EIRA cohort, together with other replication studies (Hinks et al. 2005; Lee et al. 2005; Orozco et al. 2005; Steer et al. 2005; Viken et al. 2005; Zhernakova et al. 2005), confirm the hypothesis originally described by Begovich et al. (2004) in the NARAC cohort that the *PTPN22* R620W variant influences the risk of autoantibody-positive RA. We extend the finding that *PTPN22* is associated with RF+ disease to include CCP+ disease. This result is not surprising, given the strong correlation between RF and CCP status. However, it is interesting to note that a more significant result was obtained with CCP status (table 3), which is consistent with the clinical observation that CCP is a more specific marker for RA than is RF (Schellekens et al. 2000). Together, these results provide the most convincing evidence to date of a non-MHC gene that influences the risk of developing RA.

Our stratified analyses of *PTPN22* revealed novel associations with clinically relevant subsets of RA. We identified an association of *PTPN22* with age at onset ($P = .004$), such that carriers of the susceptible allele develop RA ~2 years earlier than noncarriers do. We observed a similar result for the *HLA-DRB1* SE alleles and age at onset ($P = .0002$; data not shown). We also observed that *PTPN22* has a greater effect in RF+ males than in RF+ females ($P = .03$), a result which is no longer significant at the $P < .05$ level when we correct for our three exploratory hypotheses. Recently, Orozco et al. (2005) reported a similar trend in male versus female patients with RA. Limited clinical data for RA suggest that sex differences indeed exist (Weyand et al. 1998; Kuiper et al. 2001; Tengstrand et al. 2004), findings corroborated by animal models of arthritis (Remmers et al. 2002; Adarichev et al. 2003; Meng et al. 2004). These novel findings require replication in independent sample collections.

Interpreting the significance of *CTLA4* and *PADI4* is more challenging, but our data support the hypothesis that genetic variation within these genes influences the risk of developing RA. Our results are significant only in the NARAC cohort, although a trend toward an association with *CTLA4* is noted in the EIRA cohort. Because the cohorts are clinically different, we stratified our analysis by sex, autoantibody status, and age at onset. In support of a true-positive association for the CT60 allele of *CTLA4,* our results are strengthened in the EIRA cohort when only CCP+ cases are compared with all controls (combined OR 1.18; $P = .0006$) or with CCP− cases (OR 1.18; $P = .007$). This result, if replicated, may indicate that *CTLA4* influences the development of RA only in CCP+ patients. Our results remain significant after a formal meta-analysis with the published RA literature that includes individuals of European ancestry. Recently, Lei et al. (2005) demonstrated an association between the CT60 allele and RA susceptibility in Han Chinese (OR 1.41; 95% CI 1.10–

Table 7

Results Testing Our Exploratory Hypothesis That *PTNP22* Has a Greater Effect on RA Susceptibility in Males Than in Females

	Males (n = 1,073)					Females (n = 3,025)				
Collection	Allele T	Allele C	MAF	OR (95% CI)	P	Allele T	Allele C	MAF	OR (95% CI)	P
EIRA:				1.90 (1.25–2.88)	.003				1.20 (.92–1.56)	.176
RF+ cases	111	439	.20			210	1,126	.16		
RF− cases	32	240	.12			94	604	.13		
NARAC:				1.17 (.59–2.31)	.65				1.06 (.80–1.41)	.694
RF+ cases	75	425	.15			238	1,216	.16		
RF− cases	11	73	.13			73	395	.16		
Arthritis cohort 3:				1.96 (1.18–3.25	.01				1.34 (1.01–1.79)	.05
RF+ cases	83	433	.16			197	1,231	.14		
RF− cases	20	204	.09			71	595	.11		
Pooled				1.76 (1.31–2.36)	.00009				1.19 (1.02–1.40)	.015

Note.—We performed a case-only analysis stratified on the basis of sex and RF status. Pooled analysis of the OR of all three collections demonstrated that the effect of the *PTPN22* susceptible T allele is greater in males (OR 1.76) than in females (OR 1.19), a difference that is statistically significant by logistic regression ($P = .03$).

Table 8

Results Testing Our Exploratory Hypothesis That *CTLA4* Is Associated with CCP+ but Not CCP− Disease in Both the EIRA and NARAC Cohorts

	Allele			Cases vs. All Controls		CCP+ vs. CCP− Cases	
Collection	G	A	Frequency	OR (95% CI)	*P*	OR (95% CI)	*P*
EIRA:						1.18 (1.02–1.38)	.03
CCP+ cases	1,152	656	.64	1.13 (.98–1.29)	.08		
CCP− cases	694	468	.60	.95 (.82–1.11)	.51		
NARAC:						1.15 (.90–1.48)	.26
CCP+ cases	697	447	.61	1.26 (1.08–1.47)	.003		
CCP− cases	185	137	.57	1.09 (.86–1.39)	.48		
Pooled:						1.18 (1.03–1.34)	.007
CCP+ cases	…	…	…	1.18 (1.07–1.31)	.0006		
CCP− cases	…	…	…	.99 (.87–1.12)	.57		

Note.—The allele-frequency distribution was compared among in CCP+ cases, CCP− cases, and all controls. Only CCP+ patients demonstrated a statistically significant difference in allele frequency when compared with all controls (pooled $P = .0006$). The difference between CCP+ and CCP− cases is statistically significant ($P = .007$).

1.82; $P = .005$), providing additional support for a true-positive association.

An alternate explanation for the modest association of *CTLA4* and *PADI4* with RA in our collection is that our result is a false-positive. Possible explanations for a false-positive result include technical artifact, population stratification, or an inappropriate statistical threshold of significance. Our genotyping efficiency and accuracy argue strongly against technical artifact (see the "Material and Methods" section). Both cohorts were matched on the basis of ethnicity, which reduces the likelihood of stratification. However, it is becoming increasingly clear that subtle stratification exists even in well-matched cohorts (Freedman et al. 2004; Marchini et al. 2004). Extensive genotyping of genomic controls (Pritchard and Donnelly 2001), which is beyond the scope of this project, is necessary to assess the extent of stratification in each cohort. Ultimately, additional studies in sufficiently powered cohorts of patients with RA, together with an updated meta-analysis of the published literature, are required to determine definitively whether genetic variation within *PADI4* and/or *CTLA4* influences the risk of developing RA.

If our *CTLA4* and *PADI4* results in >4,000 samples are true positives, and the actual allelic OR is ~1.10 (as indicated by the pooled analysis), then it is not surprising that smaller studies have failed to detect a statistically significant result. An allele at ~40% population frequency with an OR of ~1.10 under a multiplicative model requires ~2,700 cases and ~2,700 controls to have 80% power to detect a true-positive result (type I error rate 0.05) (see Genetic Power Calculator Web site), and previous studies were conducted with fewer samples than this (Barton et al. 2004*a*, 2004*b*; Orozco et al. 2004; Martinez et al. 2005). Consistent with this explanation, previous studies have shown that the putative susceptibility alleles were marginally increased in patients with RA, but not to a level that reached statistical significance (see table 10 for *CTLA4* meta-analysis).

Our finding that the majority of previous association studies in RA may represent false-positive results is consistent with the published literature of other complex genetic diseases (Ioannidis et al. 2001; Lohmueller et al. 2003). Although it is possible that some of our results are false negatives, we have >80% power to detect the published OR in 14 of the 17 alleles selected for replication (and >50% for all alleles). We selected a nominal *P* value of .05 as significant, to minimize false-positive results and to maximize our confidence in a true-positive association. On the basis of empiric and theoretical data in complex traits, others have advocated similar replication *P* values (Lohmueller et al. 2003; Wacholder et al. 2004). Furthermore, a formal meta-analysis failed to detect a subtle effect that may have been missed by our replication in the NARAC and EIRA cohorts (table 9). Although we observed no significant evidence of genetic heterogeneity across studies, we excluded individuals of non-European descent. Of note, there are no formal tests of heterogeneity in patient selection criteria, and it may be that an alternate analytic method is necessary to combine effects across studies.

Other explanations for the lack of replication in our collection include genetic, ethnic, and clinical heterogeneity across study populations and the possibility that we did not test the actual causal allele in a given gene.

Table 9

Meta-Analysis of *IL4, MIF, TNFRSF1B,* and *CARD15* Allele-Frequency Data

The table is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Table 10

Meta-Analysis of *CTLA4* CT60 Allele (*rs3087243*)

Collection	Allele		Frequency	OR (95% CI)	*P*
	G	A			
EIRA:				1.05 (.93–1.19)	.21
Cases	1,870	1,140	.62		
Controls	1,070	686	.61		
NARAC:				1.23 (1.08–1.42)	.001
Cases	1,003	653	.61		
Controls	934	756	.55		
Barton et al. (2004*b*):				1.04 (.90–1.20)	.64
Cases	820	618	.57		
Controls	848	662	.56		
Orozco et al. (2004):				.98 (.81–1.19)	.84
Cases	432	434	.50		
Controls	401	395	.50		
Pooled				1.08 (1.01–1.17)	.01

Note.—Allele-frequency data of RA cases versus controls were compared across four study populations, by use of a fixed-effects (Mantel-Haenszel) meta-analysis of the ORs. No significant heterogeneity was observed among the four studies. No data on CCP status were available from the two published studies to compare CCP+ cases only (for which we observed a more significant result).

It is possible that a true susceptibility allele in one ethnic population has no effect in a different ethnic population ("ethnic heterogeneity"). There is little evidence to support this hypothesis (Ioannidis et al. 2004), although a thorough evaluation is not yet possible, given the limited number of true susceptibility alleles tested across multiple ethnic populations. Differences in clinical features or ascertainment criteria across RA collections may account for the inability to replicate. For example, we observe a more significant association between RA and *CTLA4* and *PTPN22* alleles when CCP status is taken into consideration. In an attempt to control for clinical differences, we used our literature review to define secondary hypotheses that test for relevant subsets, but we found no evidence of association. Finally, we cannot exclude that other variants within these genes may be associated with RA. We selected alleles with the most-significant association from a previous study. If we assume that an allele is either the putative causal allele or in very strong LD with the causal allele, then our study design should replicate the original finding. Comprehensive testing of genetic variation within each gene may reveal alternate alleles that contribute to RA susceptibility, as recently suggested by a study of *PTPN22* (Carlton et al. 2005).

We observed a difference in the ORs between the EIRA and NARAC cohorts for *PTPN22, CTLA4,* and *PADI4*. There are two possible explanations for the differences. First, the effect in NARAC and EIRA could actually be the same, with the different ORs simply reflecting sampling variance across the two populations. This is clearly not the case for *PTPN22,* which demonstrates a statistically significant difference between the two cohorts. Second, the OR in NARAC and EIRA could indeed be different as a result of ascertainment, clinical severity, "genetic loading" of affected sibling pairs, or an unmeasured variable. Our stratified analyses failed to identify any clear explanation for the difference, although we were unable to correct for disease severity, because of the lack of a suitable clinical variable in EIRA (e.g., disease activity score or radiographic erosions). The statistically significant differences in the ORs observed for *CTLA4* and *PADI4* are sufficiently modest such that we do not believe it is possible to determine these two association models at present. Future studies aimed at identification of covariates will be important for understanding the impact of these genes on RA clinical heterogeneity.

We did observe a P value $<.05$ for *HACVR1* and *SUMO4,* but not with the published susceptibility allele. Therefore, these results may represent statistical fluctuation rather than a true-positive association. Recent data have called into question the validity of the *SUMO4* association with type 1 diabetes (Park et al. 2005; Qu et al. 2005; Smyth et al. 2005). It should also be noted that the allele frequency of the *SLC22A4* variant tested is significantly lower in our collection (minor-allele frequency [MAF] ≈ 0.08) than in the Japanese population (MAF ≈ 0.31), which limits our power to detect a true association (especially under a recessive genetic model). Studies in additional populations are necessary to test whether these or alternate alleles are associated with RA.

In conclusion, our results replicate the finding that *PTPN22* is associated with RF+ RA and suggest the novel hypotheses that the allele influences age at RA

onset and has a greater effect in males than in females. In addition, our study provides support for association of RA with *CTLA4* and *PADI4,* a result that will need to be confirmed in additional studies. We find no evidence to support an association for the remaining alleles, which suggests that these alleles are not associated with RA with an OR $\geq$1.10. These results add to our understanding of the genetic basis of RA and emphasize the need for large collections of patients, updated meta-analysis, and stringent interpretation of statistical significance in future RA candidate-gene association studies.

Acknowledgments

R.M.P. is supported for this work by the National Institutes of Health (NIH) (grant K08 AI 55314). P.K.G. is supported for this work by the NIH (grants NO1-AR-2-2263 and R01-AR44222) and the National Arthritis Foundation. J.D.R. is supported by the National Institute of Diabetes and Digestive and Kidney Diseases and by the Crohn's and Colitis Foundation of America. This research was supported in part by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the NIH. E.W.K. is supported by NIH grants R01 AR49880-01, K24 AR0524-01, P60 AR47782-05. We are grateful to Abbott Laboratories, for the support of the SONORA study, and to Clair Bombardier, Elena Massarotti, and Michael Weisman, for their critical roles in the design and implementation of the SONORA cohort. We appreciate Julie Le's technical contribution to genotyping done at the National Institute of Arthritis and Musculoskeletal and Skin Diseases and Noel Burtt and Christine Curley's technical assistance at the Broad Institute.

Web Resources

The URLs for data presented herein are as follows:

Genetic Association Database, http://geneticassociationdb.nih.gov/

Genetic Association Models, http://statgen.iop.kcl.ac.uk/gpc/model/

Genetic Power Calculator, http://statgen.iop.kcl.ac.uk/gpc/

New York Cancer Project (NYPC), http://www.amdec.org/amdec_initiatives/nycp.html

North American Rheumatoid Arthritis Consortium (NARAC), http://www.naracdata.org/

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for RA, *HLA-DRB1, PTPN22, PADI4,* and *CTLA4*)

S.O.N.O.R.A.—Study of New Onset Rheumatoid Arthritis, http://www.abbottimmunology.com/ri_sonora.asp

References

Adarichev VA, Nesterovitch AB, Bardos T, Biesczat D, Chandrasekaran R, Vermes C, Mikecz K, Finnegan A, Glant TT (2003) Sex effect on clinical and immunologic quantitative trait loci in a murine model of rheumatoid arthritis. Arthritis Rheum 48:1708–1720

Addo A, Le J, Li W, Aksentijevich I, Balow JJ, Lee A, Gregersen P, Kastner D, Remmers E (2005) Analysis of CARD15/NOD2 haplotypes fails to identify common variants associated with rheumatoid arthritis susceptibility. Scand J Rheumatol 34:198–203

Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, Healey LA, et al (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 31:315–324

Barton A, Bowes J, Eyre S, Spreckley K, Hinks A, John S, Worthington J (2004*a*) A functional haplotype of the *PADI4* gene associated with rheumatoid arthritis in a Japanese population is not associated in a United Kingdom population. Arthritis Rheum 50:1117–1121

Barton A, John S, Ollier WE, Silman A, Worthington J (2001) Association between rheumatoid arthritis and polymorphism of tumor necrosis factor receptor II, but not tumor necrosis factor receptor I, in Caucasians. Arthritis Rheum 44:61–65

Barton A, Jury F, Eyre S, Bowes J, Hinks A, Ward D, Worthington J (2004*b*) Haplotype analysis in simplex families and novel analytic approaches in a case-control cohort reveal no evidence of association of the CTLA-4 gene with rheumatoid arthritis. Arthritis Rheum 50:748–752

Barton A, Lamb R, Symmons D, Silman A, Thomson W, Worthington J, Donn R (2003) Macrophage migration inhibitory factor (MIF) gene polymorphism is associated with susceptibility to but not severity of inflammatory polyarthritis. Genes Immun 4:487–491

Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC, Ardlie KG, et al (2004) A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (*PTPN22*) is associated with rheumatoid arthritis. Am J Hum Genet 75:330–337

Bombardier C, Deaton R, Gregersen PK, Massarotti EM, Formica C, Weisman M (2002) Pattern of DMARD use in a North American cohort of patients with early rheumatoid arthritis (SONORA) [abstract]. Arthritis Rheum 46:S344

Cantagrel A, Navaux F, Loubet-Lescoulie P, Nourhashemi F, Enault G, Abbal M, Constantin A, Laroche M, Mazieres B (1999) Interleukin-1β, interleukin-1 receptor antagonist, interleukin-4, and interleukin-10 gene polymorphisms: relationship to occurrence and severity of rheumatoid arthritis. Arthritis Rheum 42:1093–1100

Carlton VE, Hu X, Chokkalingam AP, Schrodi SJ, Brandon R, Alexander HC, Chang M, Catanese JJ, Leong DU, Ardlie KG, Kastner DL, Seldin MF, Criswell LA, Gregersen PK, Beasley E, Thomson G, Amos CI, Begovich AB (2005) *PTPN22* genetic variation: evidence for multiple variants associated with rheumatoid arthritis. Am J Hum Genet 77:567–581

Chae SC, Song JH, Shim SC, Yoon KS, Chung HT (2004) The exon 4 variations of *Tim-1* gene are associated with rheumatoid arthritis in a Korean population. Biochem Biophys Res Commun 315:971–975

Dahlqvist SR, Arlestig L, Sikstrom C, Linghult S (2002) Tumor necrosis factor receptor type II (exon 6) and interleukin-6 (-174) gene polymorphisms are not associated with family history but tumor necrosis factor receptor type II is associated with hypertension in patients with rheumatoid arthritis from northern Sweden. Arthritis Rheum 46:3096–3098

Daly MJ, Rioux JD (2004) New approaches to gene hunting in IBD. Inflamm Bowel Dis 10:312–317

Deighton CM, Walker DJ, Griffiths ID, Roberts DF (1989) The contribution of HLA to rheumatoid arthritis. Clin Genet 36:178–182

Dieude P, Petit E, Cailleau-Moindrault S, Osorio J, Pierlot C, Martinez M, Faure S, Alibert O, Lasbleiz S, De Toma C, Bardin T, Prum B, Cornelis F (2002) Association between tumor necrosis factor receptor II and familial, but not sporadic, rheumatoid arthritis: evidence for genetic heterogeneity. Arthritis Rheum 46:2039–2044

Fabris M, Tolusso B, Di Poi E, Assaloni R, Sinigaglia L, Ferraccioli G (2002) Tumor necrosis factor-α receptor II polymorphism in patients from southern Europe with mild-moderate and severe rheumatoid arthritis. J Rheumatol 29:1847–1850

Ferreiros-Vidal I, Barros F, Pablos JL, Carracedo A, Gomez-Reino JJ, Gonzalez A (2003) CARD15/NOD2 analysis in rheumatoid arthritis susceptibility. Rheumatology (Oxford) 42:1380–1382

Firestein GS (2003) Evolving concepts of rheumatoid arthritis. Nature 423:356–361

Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, Gabriel SB, Topol EJ, Smoller JW, Pato CN, Pato MT, Petryshen TL, Kolonel LN, Lander ES, Sklar P, Henderson B, Hirschhorn JN, Altshuler D (2004) Assessing the impact of population stratification on genetic association studies. Nat Genet 36:388–393

Fries JF, Wolfe F, Apple R, Erlich H, Bugawan T, Holmes T, Bruce B (2002) HLA-DRB1 genotype associations in 793 white patients from a rheumatoid arthritis inception cohort: frequency, severity, and treatment bias. Arthritis Rheum 46: 2320–2329

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. Science 296: 2225–2229

Gabriel SE, Crowson CS, O'Fallon WM (1999) The epidemiology of rheumatoid arthritis in Rochester, Minnesota, 1955–1985. Arthritis Rheum 42:415–420

Gregersen PK, Silver J, Winchester RJ (1987) The shared epitope hypothesis: an approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum 30:1205–1213

Guo D, Li M, Zhang Y, Yang P, Eckenrode S, Hopkins D, Zheng W, et al (2004) A functional variant of SUMO4, a new IκBα modifier, is associated with type 1 diabetes. Nat Genet 36:837–841

Hall FC, Weeks DE, Camilleri JP, Williams LA, Amos N, Darke C, Gibson K, Pile K, Wordsworth BP, Jessop JD (1996) Influence of the HLA-DRB1 locus on susceptibility and severity in rheumatoid arthritis. QJM 89:821–829

Hinks A, Barton A, John S, Bruce I, Hawkins C, Griffiths CE, Donn R, Thomson W, Silman A, Worthington J (2005) Association between the *PTPN22* gene and rheumatoid arthritis and juvenile idiopathic arthritis in a UK population: further support that *PTPN22* is an autoimmunity gene. Arthritis Rheum 52:1694–1699

Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95–108

Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K (2002) A comprehensive review of genetic association studies. Genet Med 4:45–61

Huang CM, Wu MC, Wu JY, Tsai FJ (2002) No association of interleukin-4 gene polymorphisms in Chinese patients with rheumatoid arthritis in Taiwan. Clin Exp Rheumatol 20:871–872

Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, Almer S, Tysk C, O'Morain CA, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahbatou M, Thomas G (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 411: 599–603

Ioannidis JP, Ntzani EE, Trikalinos TA (2004) "Racial" differences in genetic effects for complex diseases. Nat Genet 36:1312–1318

Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG (2001) Replication validity of genetic association studies. Nat Genet 29:306–309

Jawaheer D, Li W, Graham RR, Chen W, Damle A, Xiao X, Monteiro J, Khalili H, Lee A, Lundsten R, Begovich A, Bugawan T, Erlich H, Elder JT, Criswell LA, Seldin MF, Amos CI, Behrens TW, Gregersen PK (2002) Dissecting the genetic complexity of the association between human leukocyte antigens and rheumatoid arthritis. Am J Hum Genet 71:585–594

Jawaheer D, Lum RF, Amos CI, Gregersen PK, Criswell LA (2004) Clustering of disease features within 512 multicase rheumatoid arthritis families. Arthritis Rheum 50:736–741

Jawaheer D, Seldin MF, Amos CI, Chen WV, Shigeta R, Monteiro J, Kern M, Criswell LA, Albani S, Nelson JL, Clegg DO, Pope R, Schroeder HW Jr, Bridges SL Jr, Pisetsky DS, Ward R, Kastner DL, Wilder RL, Pincus T, Callahan LF, Flemming D, Wener MH, Gregersen PK (2001) A genome-wide screen in multiplex rheumatoid arthritis families suggests genetic overlap with other autoimmune diseases. Am J Hum Genet 68:927–936

Kochi Y, Yamada R, Suzuki A, Harley JB, Shirasawa S, Sawada T, Bae SC, Tokuhiro S, Chang X, Sekine A, Takahashi A, Tsunoda T, Ohnishi Y, Kaufman KM, Kang CP, Kang C, Otsubo S, Yumura W, Mimori A, Koike T, Nakamura Y, Sasazuki T, Yamamoto K (2005) A functional variant in *FCRL3,* encoding Fc receptor-like 3, is associated with rheumatoid arthritis and several autoimmunities. Nat Genet 37: 478–485

Kroot EJ, de Jong BA, van Leeuwen MA, Swinkels H, van den

Hoogen FH, van't Hof M, van de Putte LB, van Rijswijk MH, van Venrooij WJ, van Riel PL (2000) The prognostic value of anti-cyclic citrullinated peptide antibody in patients with recent-onset rheumatoid arthritis. Arthritis Rheum 43: 1831–1835

Kuiper S, van Gestel AM, Swinkels HL, de Boo TM, da Silva JA, van Riel PL (2001) Influence of sex, age, and menopausal state on the course of early rheumatoid arthritis. J Rheumatol 28:1809–1816

Kyogoku C, Tsuchiya N, Shibue T, Tokunaga K, Matsuta K (2003) *TNFR2* position 196 polymorphism in Japanese patients with rheumatoid arthritis: comment on the article by Dieude et al. Arthritis Rheum 48:273–274

Lee AT, Li W, Liew A, Bombardier C, Weisman M, Massarotti EM, Kent J, Wolfe F, Begovich AB, Gregersen PK (2005) The PTPN22 R620W polymorphism associates with RF positive rheumatoid arthritis in a dose-dependent manner but not with HLA-SE status. Genes Immun 6:129–133

Lei C, Dongqing Z, Yeqing S, Oaks MK, Lishan C, Jianzhong J, Jie Q, Fang D, Ningli L, Xinghai H, Daming R (2005) Association of the CTLA-4 gene with rheumatoid arthritis in Chinese Han population. Eur J Hum Genet 13:823–828

Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN (2003) Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 33:177–182

Maksymowych WP, Suarez-Almazor ME, Buenviaje H, Cooper BL, Degeus C, Thompson M, Russell AS (2002) HLA and cytokine gene polymorphisms in relation to occurrence of palindromic rheumatism and its progression to rheumatoid arthritis. J Rheumatol 29:2319–2326

Marchini J, Cardon LR, Phillips MS, Donnelly P (2004) The effects of human population structure on large genetic association studies. Nat Genet 36:512–517

Marchini J, Donnelly P, Cardon LR (2005) Genome-wide strategies for detecting multiple loci that influence complex diseases. Nat Genet 37:413–417

Martinez A, Valdivia A, Pascual-Salcedo D, Lamas JR, Fernandez-Arquero M, Balsa A, Fernandez-Gutierrez B, de la Concha EG, Urcelay E (2005) PADI4 polymorphisms are not associated with rheumatoid arthritis in the Spanish population. Rheumatology (Oxford) 44:1263–1266

Meng HC, Griffiths MM, Remmers EF, Kawahito Y, Li W, Neisa R, Cannon GW, Wilder RL, Gulko PS (2004) Identification of two novel female-specific non-major histocompatibility complex loci regulating collagen-induced arthritis severity and chronicity, and evidence of epistasis. Arthritis Rheum 50:2695–2705

Miterski B, Drynda S, Boschow G, Klein W, Oppermann J, Kekow J, Epplen JT (2004) Complex genetic predisposition in adult and juvenile rheumatoid arthritis. BMC Genet 5:2

Newman B, Wintle RF, van Oene M, Yazdanpanah M, Owen J, Johnson B, Gu X, Amos CI, Keystone E, Rubin LA, Siminovitch KA (2005) *SLC22A4* polymorphisms implicated in rheumatoid arthritis and Crohn's disease are not associated with rheumatoid arthritis in a Canadian Caucasian population. Arthritis Rheum 52:425–429

Newton JL, Harney SM, Timms AE, Sims AM, Rockett K, Darke C, Wordsworth BP, Kwiatkowski D, Brown MA (2004) Dissection of class III major histocompatibility complex haplotypes associated with rheumatoid arthritis. Arthritis Rheum 50:2122–2129

Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, Britton H, Moran T, Karaliuskas R, Duerr RH, Achkar JP, Brant SR, Bayless TM, Kirschner BS, Hanauer SB, Nunez G, Cho JH (2001) A frameshift mutation in *NOD2* associated with susceptibility to Crohn's disease. Nature 411: 603–606

Olerup O, Zetterquist H (1992) HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation. Tissue Antigens 39:225–235

Orozco G, Sanchez E, Gonzalez-Gay MA, Lopez-Nevot MA, Torres B, Caliz R, Ortego-Centeno N, Jimenez-Alonso J, Pascual-Salcedo D, Balsa A, de Pablo R, Nunez-Roldan A, Gonzalez-Escribano MF, Martin J (2005) Association of a functional single-nucleotide polymorphism of *PTPN22*, encoding lymphoid protein phosphatase, with rheumatoid arthritis and systemic lupus erythematosus. Arthritis Rheum 52:219–224

Orozco G, Torres B, Nunez-Roldan A, Gonzalez-Escribano MF, Martin J (2004) Cytotoxic T-lymphocyte antigen-4-CT60 polymorphism in rheumatoid arthritis. Tissue Antigens 64:667–670

Padyukov L, Silva C, Stolt P, Alfredsson L, Klareskog L (2004) A gene-environment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. Arthritis Rheum 50:3085–3092

Park Y, Park S, Kang J, Yang S, Kim D (2005) Assessing the validity of the association between the SUMO4 M55V variant and risk of type 1 diabetes. Nat Genet 37:112

Pritchard JK, Donnelly P (2001) Case-control studies of association in structured or admixed populations. Theor Popul Biol 60:227–237

Prokunina L, Padyukov L, Bennet A, de Faire U, Wiman B, Prince J, Alfredsson L, Klareskog L, Alarcon-Riquelme M (2004) Association of the PD-1.3A allele of the *PDCD1* gene in patients with rheumatoid arthritis negative for rheumatoid factor and the shared epitope. Arthritis Rheum 50: 1770–1773

Purcell S, Cherny SS, Sham PC (2003) Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19:149–150

Qu H, Bharaj B, Liu XQ, Curtis JA, Newhook LA, Paterson AD, Hudson TJ, Polychronakos C (2005) Assessing the validity of the association between the SUMO4 M55V variant and risk of type 1 diabetes. Nat Genet 37:111–112

Remmers EF, Joe B, Griffiths MM, Dobbins DE, Dracheva SV, Hashiramoto A, Furuya T, Salstrom JL, Wang J, Gulko PS, Cannon GW, Wilder RL (2002) Modulation of multiple experimental arthritis models by collagen-induced arthritis quantitative trait loci isolated in congenic rat lines: different effects of non-major histocompatibility complex quantitative trait loci in males and females. Arthritis Rheum 46: 2225–2234

Rigby AS, Silman AJ, Voelm L, Gregory JC, Ollier WE, Khan

MA, Nepom GT, Thomson G (1991) Investigating the HLA component in rheumatoid arthritis: an additive (dominant) mode of inheritance is rejected, a recessive mode is preferred. Genet Epidemiol 8:153–175

Rioux JD, Daly MJ, Silverberg MS, Lindblad K, Steinhart H, Cohen Z, Delmonte T, et al (2001) Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. Nat Genet 29:223–228

Rodriguez MR, Nunez-Roldan A, Aguilar F, Valenzuela A, Garcia A, Gonzalez-Escribano MF (2002) Association of the CTLA4 3′ untranslated region polymorphism with the susceptibility to rheumatoid arthritis. Hum Immunol 63:76–81

Schellekens GA, Visser H, de Jong BA, van den Hoogen FH, Hazes JM, Breedveld FC, van Venrooij WJ (2000) The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum 43: 155–163

Smyth DJ, Howson JM, Lowe CE, Walker NM, Lam AC, Nutland S, Hutchings J, Tuomilehto-Wolf E, Tuomilehto J, Guja C, Ionescu-Tirgoviste C, Undlien DE, Ronningen KS, Savage D, Dunger DB, Twells RC, McArdle WL, Strachan DP, Todd JA (2005) Assessing the validity of the association between the SUMO4 M55V variant and risk of type 1 diabetes. Nat Genet 37:110–111

Stastny P, Fink CW (1977) HLA-Dw4 in adult and juvenile rheumatoid arthritis. Transplant Proc 9:1863–1866

Steer S, Fisher SA, Fife M, Cuthbert A, Newton J, Wordsworth P, Lewis CM, Mathew CG, Lanchbury JS (2003) Development of rheumatoid arthritis is not associated with two polymorphisms in the Crohn's disease gene *CARD15*. Rheumatology (Oxford) 42:304–307

Steer S, Lad B, Grumley JA, Kingsley GH, Fisher SA (2005) Association of R602W in a protein tyrosine phosphatase gene with a high risk of rheumatoid arthritis in a British population: evidence for an early onset/disease severity effect. Arthritis Rheum 52:358–360

Stoll M, Corneliussen B, Costello CM, Waetzig GH, Mellgard B, Koch WA, Rosenstiel P, Albrecht M, Croucher PJ, Seegert D, Nikolaus S, Hampe J, Lengauer T, Pierrou S, Foelsch UR, Mathew CG, Lagerstrom-Fermer M, Schreiber S (2004) Genetic variation in *DLG5* is associated with inflammatory bowel disease. Nat Genet 36:476–480

Stolt P, Bengtsson C, Nordmark B, Lindblad S, Lundberg I, Klareskog L, Alfredsson L (2003) Quantification of the influence of cigarette smoking on rheumatoid arthritis: results from a population based case-control study, using incident cases. Ann Rheum Dis 62:835–841

Suzuki A, Yamada R, Chang X, Tokuhiro S, Sawada T, Suzuki M, Nagasaki M, Nakayama-Hamada M, Kawaida R, Ono M, Ohtsuki M, Furukawa H, Yoshino S, Yukioka M, Tohma S, Matsubara T, Wakitani S, Teshima R, Nishioka Y, Sekine A, Iida A, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K (2003) Functional haplotypes of *PADI4,* encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 34:395–402

Swanberg M, Lidman O, Padyukov L, Eriksson P, Akesson E, Jagodic M, Lobell A, Khademi M, Borjesson O, Lindgren CM, Lundman P, Brookes AJ, Kere J, Luthman H, Alfredsson L, Hillert J, Klareskog L, Hamsten A, Piehl F, Olsson T (2005) *MHC2TA* is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. Nat Genet 37:486–494

Tang K, Fu DJ, Julien D, Braun A, Cantor CR, Koster H (1999) Chip-based genotyping by mass spectrometry. Proc Natl Acad Sci USA 96:10016–10020

Tengstrand B, Ahlmen M, Hafstrom I (2004) The influence of sex on rheumatoid arthritis: a prospective study of onset and outcome after 2 years. J Rheumatol 31:214–222

Tokuhiro S, Yamada R, Chang X, Suzuki A, Kochi Y, Sawada T, Suzuki M, Nagasaki M, Ohtsuki M, Ono M, Furukawa H, Nagashima M, Yoshino S, Mabuchi A, Sekine A, Saito S, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K (2003) An intronic SNP in a RUNX1 binding site of *SLC22A4*, encoding an organic cation transporter, is associated with rheumatoid arthritis. Nat Genet 35:341–348

Ueda H, Howson JM, Esposito L, Heward J, Snook H, Chamberlain G, Rainbow DB, et al (2003) Association of the T-cell regulatory gene *CTLA4* with susceptibility to autoimmune disease. Nature 423:506–511

Vaidya B, Pearce SH, Charlton S, Marshall N, Rowan AD, Griffiths ID, Kendall-Taylor P, Cawston TE, Young-Min S (2002) An association between the *CTLA4* exon 1 polymorphism and early rheumatoid arthritis with autoimmune endocrinopathies. Rheumatology (Oxford) 41:180–183

van der Helm-van Mil AHM, Dieudé P, Schonkeren JJM, Cornélis F, Huizinga TWJ (2004) No association between tumour necrosis factor receptor type 2 gene polymorphism and rheumatoid arthritis severity: a comparison of the extremes of phenotypes. Rheumatology (Oxford) 43:1232–1234

van Jaarsveld CH, ter Borg EJ, Jacobs JW, Schellekens GA, Gmelig-Meyling FH, van Booma-Frankfort C, de Jong BA, van Venrooij WJ, Bijlsma JW (1999) The prognostic value of the antiperinuclear factor, anti-citrullinated peptide antibodies and rheumatoid factor in early rheumatoid arthritis. Clin Exp Rheumatol 17:689–697

van Zeben D, Hazes JM, Zwinderman AH, Cats A, van der Voort EA, Breedveld FC (1992) Clinical significance of rheumatoid factors in early rheumatoid arthritis: results of a follow up study. Ann Rheum Dis 51:1029–1035

Viken MK, Amundsen SS, Kvien TK, Boberg KM, Gilboe IM, Lilleby V, Sollid LM, Forre OT, Thorsby E, Smerdel A, Lie BA (2005) Association analysis of the 1858C>T polymorphism in the *PTPN22* gene in juvenile idiopathic arthritis and other autoimmune diseases. Genes Immun 6:271–273

Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N (2004) Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 96:434–442

Wang WY, Barratt BJ, Clayton DG, Todd JA (2005) Genome-wide association studies: theoretical and practical concerns. Nat Rev Genet 6:109–118

Weyand CM, Schmidt D, Wagner U, Goronzy JJ (1998) The influence of sex on the phenotype of rheumatoid arthritis. Arthritis Rheum 41:817–822

Williams RL (2000) A note on robust variance estimation for cluster-correlated data. Biometrics 56:645–646

Wolfe F, Michaud K, Gefeller O, Choi HK (2003) Predicting

mortality in patients with rheumatoid arthritis. Arthritis Rheum 48:1530–1542

Yamada R, Tanaka T, Unoki M, Nagai T, Sawada T, Ohnishi Y, Tsunoda T, Yukioka M, Maeda A, Suzuki K, Tateishi H, Ochi T, Nakamura Y, Yamamoto K (2001) Association between a single-nucleotide polymorphism in the promoter of the human interleukin-3 gene and rheumatoid arthritis in Japanese patients, and maximum-likelihood estimation of combinatorial effect that two genetic loci have on susceptibility to the disease. Am J Hum Genet 68:674–685

Yamada R, Ymamoto K (2005) Recent findings on genes associated with inflammatory disease. Mutat Res 573:136–151

Yen JH, Tsai WC, Chen CJ, Ou TT, Lin CH, Hu CJ, Liu HW (2003) Tumor necrosis factor receptor 2 microsatellite and exon 6 polymorphisms in rheumatoid arthritis in Taiwan. J Rheumatol 30:438–442

Zhernakova A, Eerligh P, Wijmenga C, Barrera P, Roep BO, Koeleman BP (2005) Differential association of the PTPN22 coding variant with autoimmune diseases in a Dutch population. Genes Immun 6:459–461